



# Generative optical modeling of whole blood for detecting platelets in lens-free images

BENJAMIN D. HAEFFELE,<sup>1,\*</sup> CHRISTIAN PICK,<sup>1</sup> ZIDUO LIN,<sup>2</sup>  
EVELIEN MATHIEU,<sup>1</sup> STUART C. RAY,<sup>1</sup> AND RENÉ VIDAL<sup>1</sup>

<sup>1</sup> Johns Hopkins University, Baltimore, MD 21218, USA

<sup>2</sup> Imec, Leuven B-3001, Belgium

\*bhaeffele@jhu.edu

**Abstract:** In this paper, we consider the task of detecting platelets in images of diluted whole blood taken with a lens-free microscope. Despite having several advantages over traditional microscopes, lens-free imaging systems have the significant challenge that the resolution of the system is typically limited by the pixel dimensions of the image sensor. As a result of this limited resolution, detecting platelets is very difficult even by manual inspection of the images due to the fact that platelets occupy just a few pixels of the reconstructed image. To address this challenge, we develop an optical model of diluted whole blood to generate physically realistic simulated holograms suitable for training machine learning models in a supervised manner. We then use this model to train a convolutional neural network (CNN) for platelet detection and validate our approach by developing a novel optical configuration which allows collecting both lens-free and fluorescent microscopy images of the same field of view of diluted whole blood samples with fluorescently labeled platelets.

© 2020 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

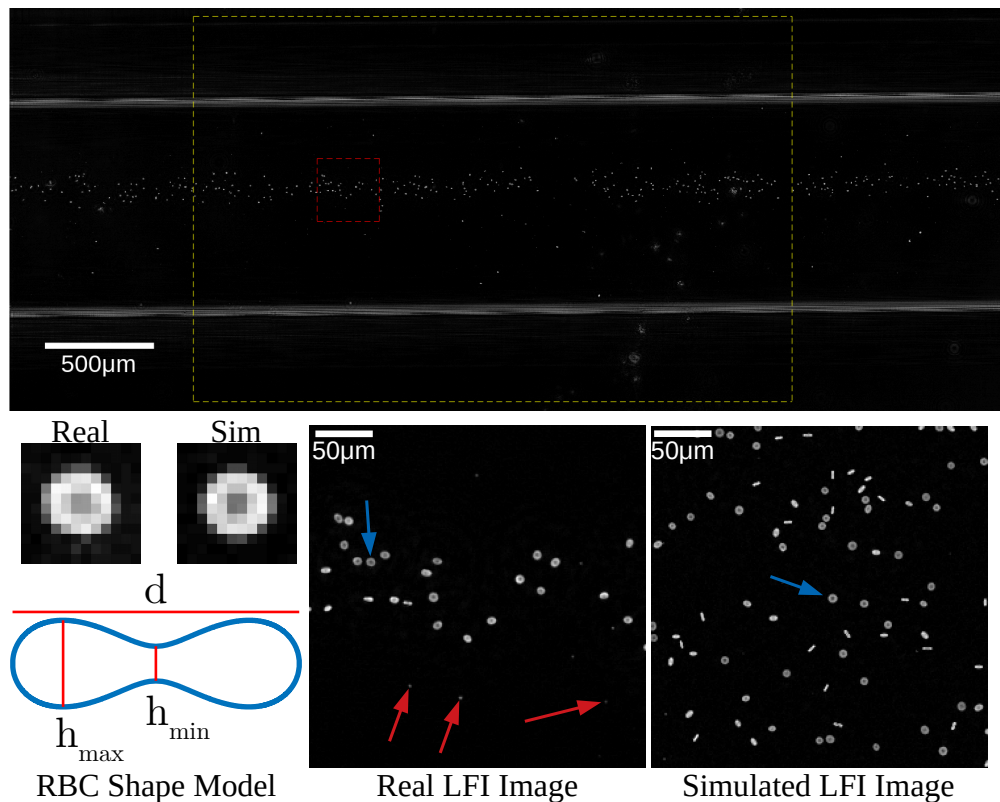
## 1. Introduction

Lens-free imaging (LFI) is a form of digital microscopic holography which records the diffraction patterns (also referred to as holograms) of a specimen illuminated with coherent light (e.g., from a laser) and then reconstructs an image of the specimen by inverting a mathematical model of the light diffraction process. LFI has multiple advantages over conventional microscopy. First, as the name implies, the system does not require lenses which significantly reduces the overall system cost, complexity, and size. Second, LFI systems have larger fields of view than traditional microscopes with equivalent magnification. Third, the system does not require any manual focusing as the focal depth can be adjusted via software, which additionally eliminates the strict mechanical stability requirements of lens-based systems (where the lens must be held at a precise focal distance from the image sensor) [1–3].

These advantages have led to LFI being explored as a potential solution for various applications in biomedical microscopy in resource-limited settings [4,5], such as semen [6] and hematological analysis [7–12]. Likewise, in this work we are interested in exploiting these advantages of LFI systems to develop a compact and low-cost system that is capable of measuring the concentration of platelets in human blood. Platelets are responsible for forming blood clots to stop bleeding in response to injury, and abnormal platelet levels can be both indicative and causative of a wide variety of pathologies. As a result, many clinical situations require routine monitoring of platelet counts, making them an indispensable medical tool and one of the standard analytes of a complete blood count (CBC), the most widely ordered blood test worldwide.

However, despite the previously mentioned advantages of LFI systems and the clinical necessity for monitoring platelet counts, a significant challenge in LFI applications is that the resolution of LFI systems is often limited by the pixel size (a.k.a. pixel pitch) of the image sensor (here our system is based on the system described in [13] and uses a 13 megapixel image sensor with a  $1.12\mu\text{m} \times 1.12\mu\text{m}$  pixel pitch). A platelet typically has a diameter of only 2–3 micrometers and a

volume of only 9-12 femtoliters, making identification of platelets in reconstructed LFI images very challenging even with close manual inspection of the image. As an example, Fig. 1 (Bottom Middle) shows a small crop from a reconstructed image of diluted whole blood. Note that while some platelets are visible (a few are denoted with red arrows), they are hard to identify manually and can be easily confused with artifacts in the reconstructed image. One potential solution is to consider digital holography techniques beyond LFI, which also allow for software focusing as well as higher image resolution, and such techniques have been used for a variety of biomedical applications such as assessing red blood cell morphology, quantifying cardiomyocyte activity, and detecting malaria (see [14,15] for reviews). In order to achieve higher resolution, however, these digital holography techniques often rely on additional optical components such as mirrors, beam-splitters, and lenses, which can significantly increase system cost and complexity and often results in a much smaller field of view than is possible with LFI systems.



**Fig. 1.** (Top) Full size reconstructed LFI image: yellow dashed rectangle - field of view that is overlapping with the fluorescent microscope described in §4, red dashed square - field of view of the crop shown in the bottom row. The horizontal lines in the image that run the full width of the image are the walls of the microfluidic channel. (Bottom Left) Parameters of the RBC shape model and crops of real and simulated RBCs. (Bottom Middle) Small crop of the example reconstructed LFI image (from red dashed square). Red arrows denote example platelets and blue arrow denotes the RBC shown in the crop to the left. (Bottom Right) Small crop from a simulated image reconstruction.

In terms of object detection methods, current state-of-the-art systems largely come from natural image applications and are predominately based on deep neural networks (see [16–18] for a few current, well-known examples). However, training large network models (or supervised

learning methods in general) requires access to significant volumes of training images along with corresponding ground-truth regarding the location of the various objects of interest. As described above, it is often very challenging to accurately locate and identify platelets even by manual inspection in reconstructed lens-free images due to 1) their small size relative to the resolution of the image 2) the relatively small signal that they generate relative to the other cells in the image (predominately red blood cells, which have roughly an order of magnitude larger volume), and 3) the fact that there is roughly an order of magnitude fewer platelets compared to red blood cells in a typical sample. Combined, these issues significantly limit the potential of constructing even moderately sized training and testing datasets, which precludes the use not only of large-scale supervised methods like neural networks but *any* supervised learning method (such as training a simple classifier based on pre-defined image features) as well as making the quantitative evaluation of *any* object-detection approach very challenging.

In this work, we address these challenges through two key contributions. First, we develop an optical model which allows us to simulate synthetic holograms of diluted whole blood with sufficient realism to train a convolutional neural network (CNN) capable of detecting platelets in real LFI images (which by extension also enables the use of other supervised learning methods). Additionally, we also develop a validation method for our approach by constructing a novel tandem microscopy imaging setup which allows us to record a LFI hologram and a fluorescent image of an overlapping field of view within a few seconds of each other. By fluorescently labeling platelets we then compare the platelet detections from our trained neural network operating on LFI images with detections from the corresponding fluorescent image (which is easier due to the fluorescent labeling), using detections from the fluorescent image as a presumed ground truth. This paper extends a preliminary conference publication [19] by adding additional description of the optical setup and image analysis pipeline used for the verification protocol as well as testing the method on a larger set of images.

## 2. Optical model

To develop our optical model of diluted whole blood, we need a means to optically model the various cell types present in human blood: red blood cells (RBCs), platelets (PLTs), and white blood cells (WBCs). WBCs are relatively uncommon (roughly 3 orders of magnitude lower concentrations than RBCs), so their presence or absence in an image has little impact on detecting PLTs. As a result, we will largely focus on modeling RBCs and PLTs.

### 2.1. Red blood cell model

To model the optical properties of an RBC, we use a phase-plate model which describes the modulation of the incident light wave that is created by the RBC as a phase shift proportional to the integration of the RBC shape along the optical axis (by common convention we'll use the  $z$  axis as the direction of light propagation). This model is consistent with scattering measurements taken of RBCs which have also noted that at the wavelength of light used by our LFI system (637nm) RBCs do not absorb light [20]. With this model, the optical modulation of the incident wavefront is entirely determined by the RBC shape and orientation (and more specifically the integral of the shape along the optical axis), so to model the RBC shape, we use the parametric model for RBC shape given in [21], which takes the general form:

$$(x^2 + y^2 + z^2)^2 + P(x^2 + y^2) + Qz^2 + R \leq 0, \quad (1)$$

where  $(P, Q, R)$  are coefficients determined by the minimum thickness of the RBC,  $h_{min}$ , the maximal thickness of the RBC,  $h_{max}$ , and the diameter of the RBC,  $d$ , given as:

$$\begin{aligned}
 P &= \frac{1}{2} \left[ -d^2 + h_{max}^2 \left( \left( \frac{d}{h_{min}} \right)^2 - 1 \right) \left( 1 - \sqrt{1 - \left( \frac{h_{min}}{h_{max}} \right)^2} \right) \right] \\
 R &= -\left( \frac{d}{2} \right)^2 P - \left( \frac{d}{2} \right)^4 \quad Q = \left( \frac{d}{h_{min}} \right)^2 P + \left( \frac{h_{min}}{2} \right)^2 \left( \left( \frac{d}{h_{min}} \right)^4 - 1 \right).
 \end{aligned} \tag{2}$$

Given the general shape of an RBC (see Fig. 1 (Bottom Left) for an illustration of the relevant dimensions), we then generate RBCs at arbitrary orientations and locations in the image by rotating and translating the coordinate system in (1) for each RBC and then integrate along the optical axis to produce the total path length image (recall path length is proportional to phase shift) induced by the  $k^{th}$  RBC as:

$$\bar{\theta}_k^{RBC}(x, y) = \int \mathbb{I}[(\bar{x}_k^2 + \bar{y}_k^2 + \bar{z}_k^2)^2 + P_k(\bar{x}_k^2 + \bar{y}_k^2) + Q_k \bar{z}_k^2 + R_k \leq 0] dz, \tag{3}$$

where  $\mathbb{I}[c]$  is an indicator function which takes value 1 if condition  $c$  is true and 0 otherwise,  $(\bar{x}_k, \bar{y}_k, \bar{z}_k)$  are the rotated and translated coordinates for the  $k^{th}$  RBC, and  $(P_k, Q_k, R_k)$  are the coefficients for RBC  $k$  with a prescribed set of shape dimensions  $(d_k, h_{max,k}, h_{min,k})$ , each independently sampled for each cell uniformly over the ranges  $([7, 8.5], [2, 2.5], [0.8, 1.4])\mu\text{m}$ .

## 2.2. Platelet model

Since platelets are very small relative to the resolution of images we are simulating, we use a relatively simple model in our simulations as additional details will be largely irrelevant after image discretization. Namely, we again assume that platelets modulate the light wavefront largely by simply shifting the phase of the wavefront (i.e., they do not significantly absorb light). As a result, we model platelets as being a simple disk of shifted phase

$$\theta_j^{PLT}(x, y) = \psi_j \mathbb{I}[\bar{x}_j^2 + \bar{y}_j^2 \leq r_j^2], \tag{4}$$

where  $\psi_j$  denotes a constant phase shift (which we uniformly sample from  $[0.25, 0.75]/(2\pi)$  radians based on typical measurements of platelets in our reconstructed images) applied to all pixels within platelet  $j$ , where  $(\bar{x}_j, \bar{y}_j)$  are the translated coordinates for the  $j^{th}$  platelet, and  $r_j$  is the radius of the  $j^{th}$  platelet, sampled uniformly over the range  $[0.8, 1.5]\mu\text{m}$ .

## 2.3. Full model

Given these individual models for both PLTs and RBCs, we then simulate the (complex valued) optical wavefront at the image plane (on a discrete  $1024 \times 1024$  grid of pixels with dimensions  $1.12\mu\text{m} \times 1.12\mu\text{m}$ ) by combining the various phase shifts induced by all the simulated cells of various sizes, locations, and orientations:

$$I(x, y) = \exp \left\{ 2\pi i \left[ \frac{\nu_{RBC} - \nu_{media}}{\lambda} \sum_k \bar{\theta}_k^{RBC}(x, y) + \sum_j \theta_j^{PLT}(x, y) \right] \right\}, \tag{5}$$

where  $\nu_{RBC} = 1.4$  is the refractive index of an RBC at our illumination wavelength,  $\lambda = 637\text{nm}$ , as measured in [20], and  $\nu_{media} = 1.33$  is the refractive index of the fluid media suspending the blood cells. We sample the locations of PLTs and RBCs uniformly over the entire field-of-view, and we sample the cell density (cells/image area) and percentage of cells that are PLTs (with the rest being RBCs) uniformly over the ranges  $[0.285, 3.58]$  cells/kilopixels and  $[3\%, 10\%]$ , respectively, to correspond with the typical dilutions and biological values in our experiments.

Given the simulated wavefront at the image plane,  $I(x, y)$ , we then simulate the hologram by projecting the wavefront to the image sensor plane a distance  $z_0$  away (we sample uniformly over  $[400, 1200]\mu\text{m}$  in our simulations) using the wide-angular spectrum (WAS) model for

light propagation [1] which projects the wavefront via a convolution with a transfer function  $I_{z_0}(x, y) = t_{z_0}(x, y) * I(x, y)$ , with the transfer function,  $t_{z_0}(x, y)$ , defined in Fourier space as,

$$\mathcal{F}\{t_{z_0}(x, y)\}[k_x, k_y] = \exp\left(iz_0\sqrt{\left(\frac{2\pi}{\lambda}\right)^2 - k_x^2 - k_y^2}\right). \quad (6)$$

Once the simulated wavefront is projected to the image sensor plane, we then produce a final simulated hologram by taking the absolute value of the wavefront due to the physics of the image sensor only being able to record the magnitude of the optical wavefront but not the phase. Finally, we add a small amount of sampling noise, to produce the final simulated hologram as follows,

$$H(x, y) = |I_{z_0}(x, y)| + \epsilon(x, y) \text{ where } \epsilon(x, y) \stackrel{i.i.d.}{\sim} \mathcal{N}(0, \sigma) \forall(x, y). \quad (7)$$

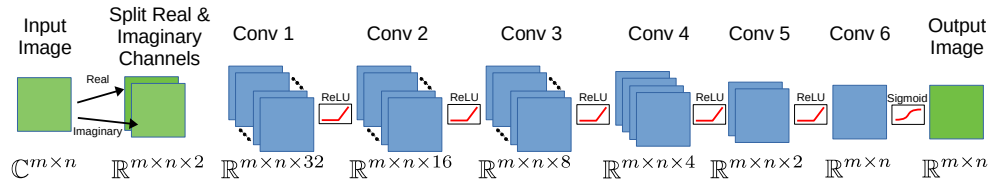
Here we have used a Gaussian noise model for the image sensor, with standard deviations uniform over the range [0.0125, 0.03125], but other noise models (e.g., Poisson) could also be employed depending on the application.

### 3. Platelet detection

Using the previously described method for simulating LFI holograms, we trained a convolutional neural network (CNN) to detect platelet locations from the recorded hologram. The first step in this process is to reconstruct an image of the specimen from the simulated hologram, for which we employ the sparse phase recovery reconstruction method developed in [22]. Figure 1 shows example reconstructions from both a real hologram and a simulated hologram, which have strong qualitative similarities. In addition to sensor noise added to the simulated hologram, we also add an offset (uniformly sampled over  $\pm 3\mu\text{m}$ ) to the reconstruction focal depth versus the true focal depth used to generate the simulated hologram to account for potential errors in auto-focusing that can occur when reconstructing real images. After reconstruction, the image is complex valued, representing an estimate of the image wavefront at the specimen plane (all images of reconstructions show the absolute value of the wavefront), so to train a CNN to detect platelets, we split the real and imaginary components of the reconstruction into two input channels to the network. The rest of the network is then a fully convolutional network, consisting of a sequence of six convolutional layers with kernels of spatial dimension 3x3 and the number of output channels reducing by a factor of 2 each layer ([32, 16, 8, 4, 2, 1], respectively). Rectified Linear Unit (ReLU) non-linearities are applied entry-wise after each convolution, with the exception of the final layer which applies a sigmoid non-linearity (see Fig. 2). The use of a fully convolutional network was done for two reasons. First, it allows the network to be applied to an input image of arbitrary size, and second, due to the small size of the platelets in these images we did not want to lose any spatial information regarding their location through any operation that reduces the image dimension (such as max-pooling or mean-pooling).

Note that the output of the network is an image with the same spatial dimension as the input, where the magnitude of each pixel is the probability that the pixel contains a platelet. As a result, we train the network as a pixel-wise classification problem using the cross-entropy loss applied pixel-wise comparing to whether a given pixel contains a platelet in the simulated image. The network weights are optimized using standard stochastic gradient descent with Nesterov acceleration. Mini-batches of 10 simulated images with dimension 1024x1024 are generated, and for each mini-batch 50 gradient descent steps are taken before a new mini-batch is generated. To perform inference on unseen real images, we simply threshold the output image at a value of 0.5 (recall the sigmoid non-linearity outputs a value in the range [0, 1]) and treat each connected component in the thresholded image as a platelet detection with no morphological filtering.





**Fig. 2.** Network architecture used for platelet detection. After splitting the complex-valued reconstruction into real and imaginary channels, the remaining network is fully convolutional with bias terms and ReLU non-linearities following each convolution with the exception of the final convolution which uses a sigmoid non-linearity. All kernels use a spatial dimension of  $3 \times 3$  with a stride of 1. The indicated dimensions correspond to the output dimension of the representation following the convolution of that layer (e.g., the output of the Conv 1 layer is  $m \times n \times 32$ ). Layers in blue contain trainable parameters.

## 4. Testing and validation

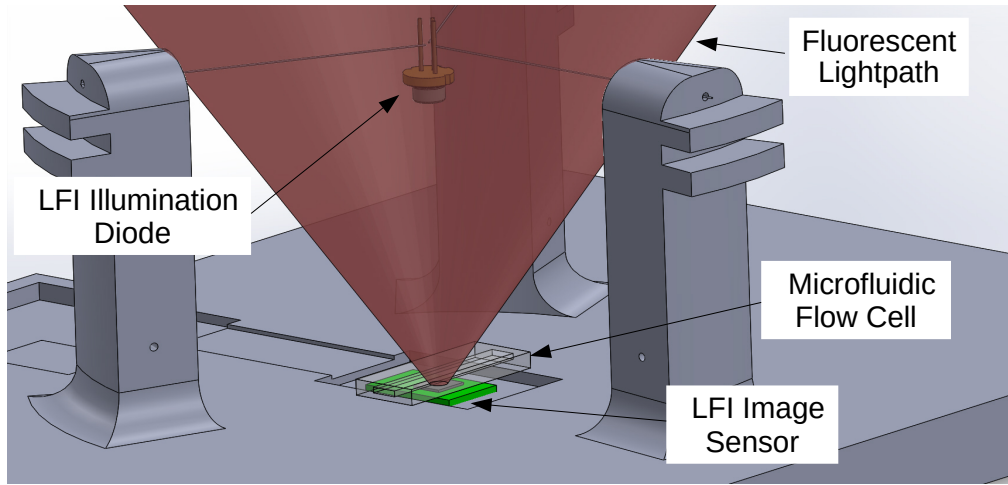
The fact that PLTs are very hard to detect even manually in LFI images presents significant challenges not only to the training of PLT detection methods (as discussed above) but also to the quantitative testing and validation of such methods as one often does not have access to high quality ground truth information. To address this issue, we developed a tandem microscopy setup which allows for both fluorescent and LFI images with a partially overlapping field of view (FOV) to be recorded within a few seconds of each other. By fluorescently labeling the PLTs we can then detect PLTs in the fluorescent images with fairly high confidence (as they are the only fluorescent objects in the image), and we then compare the set of detections in the fluorescent images with those in the LFI images.

### 4.1. Dual imaging setup

We developed a setup to allow simultaneous collection of LFI holograms paired with fluorescence microscope images in order to facilitate the positive verification of platelet detections in LFI reconstructions. The setup consists of placing the LFI optical system under an epi-fluorescence microscope (Zeiss Axiozoom v.16) with a high numerical aperture objective (0.25 NA) and long working distance (56 mm) allowing the LFI image sensor to sit directly beneath the microscope objective (see Fig. 3). The LFI laser diode is suspended by thin wires directly underneath the objective to provide illumination for the LFI image sensor while minimizing blockage of the fluorescent lightpath. While some of the fluorescent lightpath is obstructed, due to the placement of the diode, the blocked light largely corresponds to the low-frequency components of the fluorescent image, resulting in minimal distortion to the fluorescent image. Further recall that we are primarily interested in detecting the locations of the PLTs in the fluorescent images to use as a ground truth, so we can tolerate minor distortions in the fluorescent images. With this setup, images were collected by coordinating acquisition between the fluorescent microscope and LFI system, with the fluorescent microscope first collecting an image of diluted blood containing fluorescently labeled platelets followed closely by the LFI system collecting an image. A rendering of the imaging setup is given in Fig. 3.

### 4.2. Sample preparation and image acquisition

The sample itself consists of EDTA anticoagulated whole human blood containing immunolabeled platelets. Immunolabeling was accomplished by incubating samples with CD41/CD61-FITC human antibody (Miltenyi Biotec), which selectively only labels platelets. To image this label with the fluorescent microscope we illuminated the specimen with a Zeiss HXP 200C light source and used a Zeiss 38 HE filter cube, which consists of a 495nm dichroic,  $470 \pm 20$ nm excitation filter,



**Fig. 3.** Optical setup for simultaneous LFI and fluorescent imaging. The LFI illumination diode is suspended by wires inside the lightpath (in red) of the fluorescent microscope.

and a  $525 \pm 25\text{nm}$  emission filter. After labeling, the blood is diluted with phosphate-buffered saline to a final dilution of 1:200. The sample was flown through a  $50\mu\text{m}$  tall (along direction of light propagation) and  $1000\mu\text{m}$  wide (vertical axis in Fig. 1) microfluidic channel placed on top of the LFI image sensor, with a pause in sample flow during image acquisition to minimize sample movement between the LFI and fluorescent image acquisitions. The overall sequence consisted of the following steps: 1) flow of fresh sample into LFI/fluorescent microscope field of view 2) sample flow was stopped, allowing sample movement to halt before acquisition 3) fluorescent microscope acquisition 4) fluorescent illumination is shuttered 5) LFI image acquisition 6) flow of fresh sample into field of view in preparation for another pair of acquisitions. The cycle was repeated to collect a set of LFI/fluorescent microscope paired images for verification.

#### 4.3. PLT detection in fluorescent images

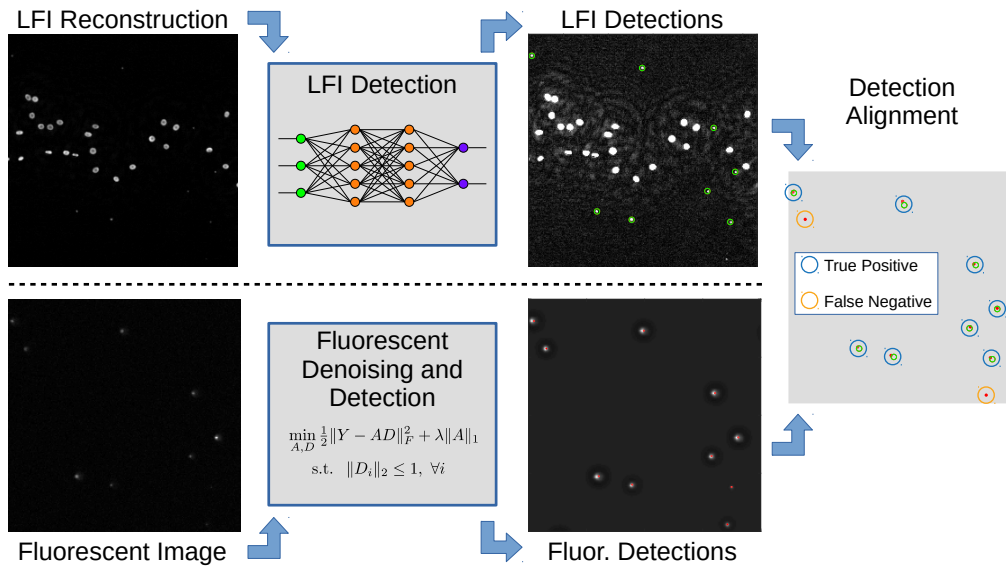
To detect the fluorescently labeled PLTs in the fluorescent images we perform a standard image denoising procedure based on sparse dictionary learning [23], where we first extract all  $10 \times 10$  pixel patches from the image using a sliding window, normalize the patches to have zero mean and unit Euclidean norm, and then train a sparse dictionary from the patches via the following formulation:

$$\min_{D,A} \frac{1}{2} \|Y - DA\|_F^2 + \lambda \|A\|_1 \text{ s.t. } \|D_i\|_2 \leq 1, \forall i \in \{1, \dots, L\}, \quad (8)$$

where  $Y \in \mathbb{R}^{100 \times p}$  is a matrix of all  $p$  extracted and normalized patches,  $D = [D_1, \dots, D_L] \in \mathbb{R}^{100 \times L}$  is the matrix of  $L$  dictionary atoms (we used  $L = 500$ ), and  $A \in \mathbb{R}^{L \times p}$  is the matrix of sparse encodings for each patch. From the learned sparse encoding, we reconstruct the patches using the sparse encoding approximation of the patches (i.e.,  $DA$ ), and finally regenerate the denoised image by returning the patches to the appropriate locations and averaging over the overlapping patches. Once the fluorescent images have been denoised via dictionary learning, the platelets are easily detected via a simple thresholding. The bottom row of Fig. 4 shows an example crop of the fluorescent image and corresponding denoised image with detections.

#### 4.4. Aligning fluorescent and LFI images

Once PLT detection has been performed on both an LFI image and a corresponding fluorescent image, we then align the coordinate systems between the two image modalities. Although the



**Fig. 4.** Pipeline for evaluating PLT detection accuracy in LFI images. (Top Row) An LFI image (Left: Example crop) is passed through a LFI detection method producing detections (Middle: Example detections from our CNN detector, grayscale saturated to better visualize platelets). (Bottom Row) A fluorescent image (Left) is denoised using the dictionary learning approach in §4.3. After denoising the platelets are detected via simple thresholding, giving the set of fluorescent detections (Middle). (Right Column) After PLTs have been detected in both the LFI and fluorescent images, then the sets of detections are aligned via the procedure described in §4.4 and the performance of the LFI detection method is measured using the metric described in §4.5.

two image modalities have a partially overlapping FOV, the two images are taken at different magnifications and spatial offsets relative to each other, so to register the two sets of image coordinates we fit an affine transformation using an alternating minimization approach where we begin with a rough estimate of the alignment transformation between the images. Then, given the assumed alignment, we project one set of PLT detections into the coordinates of the other set and match the two sets of detections using a linear assignment with an Euclidean distance cost to produce correspondences. Using the new proposed assignments we then update the parameters of the affine transformation between the coordinate systems to minimize the Euclidean error between the proposed correspondences between the two sets of detections.

#### 4.5. Performance metrics

To evaluate the performance of our combined model (a PLT detection CNN trained using images from our simulator) we collected two datasets of paired LFI and fluorescent images from two different experimental blood samples, with 39 and 99 image pairs in each dataset, respectively. We then treat the PLT detections in the fluorescent image as a ground truth and compute precision and recall scores along with the F-measure (the F-measure is defined as  $2(\text{precision} * \text{recall}) / (\text{precision} + \text{recall})$ ). Due to the fact that the LFI and fluorescent images are not collected at exactly the same time, the cells (which are suspended in a microfluidic flow-cell channel) can move slightly between image acquisitions, so even after the coordinate set alignment described above there is still some offset between detections in the two image modalities. As a result, we compute the precision, recall, and F-measure statistics as a function of an allowed detection radius, where we label a detection in the LFI image as being correct if it is within the



detection radius of a detection in the fluorescent image. Specifically, to avoid assigning multiple LFI detections to a single fluorescent detection (or vice versa), we solve a linear assignment problem between the two sets of detections ( $N_{LFI}$  LFI detections and  $N_F$  fluorescent detections) with form:

$$\min_{\Lambda} \langle \Lambda, C \rangle \quad \text{s.t.} \quad \mathbf{1}^T \Lambda \leq \mathbf{1}^T, \Lambda \mathbf{1} \leq \mathbf{1}, 0 \leq \Lambda_{ij} \leq 1 \quad \forall (i, j) \quad (9)$$

where the cost matrix  $C \in \mathbb{R}^{N_{LFI} \times N_F}$  has zero cost in matching two detections if they lie within the detection radius of each other and a cost of one if they lie outside of that radius:

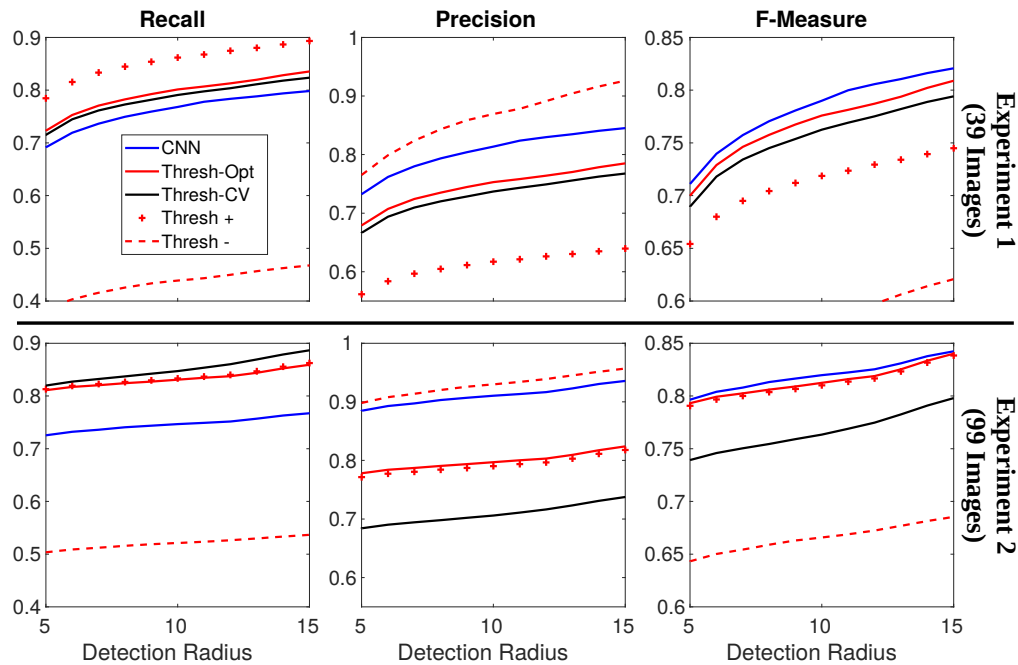
$$C_{ij} = \begin{cases} 0 & \text{if } (x_i - x_j)^2 + (y_i - y_j)^2 \leq r_d^2 \\ 1 & \text{otherwise} \end{cases} \quad (10)$$

where  $(x_i, y_i)$  are detection coordinates for detection  $i$  in the LFI image,  $(x_j, y_j)$  are detection coordinates for detection  $j$  in the fluorescent image (following coordinate alignment), and  $r_d$  is the allowed detection radius. Any detections lying outside the overlapping FOV between the two images (after coordinate alignment) were discarded prior to linear assignment.

#### 4.6. Baseline methods and results

To compare our method against a baseline, the fact that in our original problem we do not have knowledge of ground truth locations again presents difficulty as we cannot compare against any supervised approaches for object detection. As a result, we compare against several unsupervised approaches. First, we consider an “optimal” thresholding baseline, where we study the best performance that can be achieved via simple thresholding, even if the hyper-parameters of the thresholding algorithm are tuned directly to maximize performance on the test data. Specifically, we threshold the LFI image and then label a connected component in the thresholded image a PLT if its area is within lower and upper bound limits. We then maximize the F-measure at a detection radius of 15 pixels by performing an exhaustive grid search over the choice of the image intensity threshold and lower/upper bounds on the connected component area. Note that this uses full knowledge of the ground-truth to tune the thresholding hyperparameters and as a result is an over-estimate of the performance of thresholding. Figure 5 shows that even though the optimal thresholding method makes full use of the ground-truth in selecting hyperparameters our method still achieves a higher F-measure across all choices of allowed detection radii. Additionally, the optimal thresholding method is very unstable to choice of hyperparameters, and simply increasing (Thresh +) or decreasing (Thresh -) the upper and lower limits of the connected component area by the minimum increment in the grid search (3 pixels) significantly degrades performance. Further, cross validating the thresholding hyperparameters by using parameters tuned from one experiment to evaluate the other experiment results in a significant drop in F-measure performance (Thresh-CV). In contrast our CNN, trained exclusively using our developed optical simulator, achieves very good performance on both datasets with over 80% F-Measure score once the detection radius is above approximately 10 pixels, roughly on the order of the alignment error between the LFI and fluorescent coordinates due to cells drifting slightly between LFI and fluorescent image acquisitions.

In addition to the optimal thresholding baseline, we also consider the template matching (TM) baseline algorithm described in [11]. Here, we first convert the complex valued reconstruction to an absolute valued image and then use platelet templates which consist of symmetric 2D-Gaussians. We again perform an exhaustive grid search over the hyper-parameters of the TM algorithm, including the standard deviation of the Gaussian template and the threshold used in the non-maximal suppression. Even with choosing the optimal set of hyperparameters directly on test data, TM achieves very poor performance with extensive false positive triggers on RBCs, and the highest F-measure across all hyperparameter choices and allowed detection radii was less than 0.5 (off the scale of Fig. 5).



**Fig. 5.** Recall, Precision, and F-Measure values as a function of detection radius. CNN: The convolutional neural network trained using our simulated model. Thresh-Opt: Detection based on thresholding, where the optimal threshold hyper-parameters are tuned using the test data. Thresh-CV: Detection based on thresholding using the optimal threshold hyper-parameters from the other experiment (e.g., tune on Exp. 1, test on Exp. 2). Thresh +/-: The performance after making a small increase (Thresh +) or decrease (Thresh -) to the optimal connected component area hyperparameter.

## 5. Conclusions

We have presented an optical model of diluted whole blood that is sufficiently realistic to be successful in training a CNN based object detection network. Our approach achieves good performance on the very difficult task of detecting platelets in reconstructed LFI images, which can be challenging even by manual inspection due to the limited resolution of LFI systems and additionally presents significant challenges in even validating a given method. As a result, to validate our approach we developed and constructed a novel tandem microscopy setup which allows for close to simultaneous imaging of a fluorescent image (with fluorescently labeled platelets) and a LFI image of an overlapping field of view.

## Funding

National Institute on Aging (5R01AG067396-02); miDiagnostics, Inc.

## Acknowledgments

We thank Murali Jayapala, Lin Zhou, Florence Yellin, and Benjamin Bejar for useful discussions regarding this work. Funding for the study described in this publication was provided by miDiagnostics. miDiagnostics is a private company created to develop the miLab diagnostic platform. This work was also partially supported by NIH NIA grant 5R01AG067396-02.

## Disclosures

Johns Hopkins University holds a Board of Directors position in miDiagnostics, owns equity in the company, and is also eligible to receive royalty payments for commercial sales by miDiagnostics. BDH, CP, EM, SCR, and RV are eligible to receive distributions of royalty payments to JHU and/or payments related to JHU equity ownership interests in miDiagnostics. This arrangement has been reviewed and approved by Johns Hopkins University in accordance with its conflict of interest policies.

## References

1. M. K. Kim, *Digital Holographic Microscopy* (Springer, 2011).
2. S. B. Kim, H. Bae, K.-I. Koo, M. R. Dokmeci, A. Ozcan, and A. Khademhosseini, "Lens-free imaging for biological applications," *J. Lab. Autom.* **17**(1), 43–49 (2012).
3. A. Ozcan and U. Demirci, "Ultra wide-field lens-free monitoring of cells on-chip," *Lab Chip* **8**(1), 98–106 (2008).
4. S. B. Kim, H. Bae, K. Koo, M. R. Dokmeci, A. Ozcan, and A. Khademhosseini, "Lens-free imaging for biological applications," *J. Lab. Autom.* **17**(1), 43–49 (2012).
5. S. Seo, T.-W. Su, D. K. Tseng, A. Erlinger, and A. Ozcan, "Lensfree holographic imaging for on-chip cytometry and diagnostics," *Lab Chip* **9**(6), 777–787 (2009).
6. T.-W. Su, A. Erlinger, D. Tseng, and A. Ozcan, "Compact and light-weight automated semen analysis platform using lensfree on-chip microscopy," *Anal. Chem.* **82**(19), 8307–8312 (2010).
7. S. Seo, S. O. Isikman, I. Sencan, O. Mudanyali, T.-W. Su, W. Bishara, A. Erlinger, and A. Ozcan, "High-throughput lens-free blood analysis on a chip," *Anal. Chem.* **82**(11), 4621–4627 (2010).
8. H. Zhu, I. Sencan, J. Wong, S. Dimitrov, D. Tseng, K. Nagashima, and A. Ozcan, "Cost-effective and rapid blood analysis on a cell-phone," *Lab Chip* **13**(7), 1282–1288 (2013).
9. B. D. Haeffele, S. Roth, L. Zhou, and R. Vidal, "Removal of the twin image artifact in holographic lens-free imaging by sparse dictionary learning and coding," in *IEEE International Symposium on Biomedical Imaging*, (2017), pp. 741–744.
10. F. Yellin, B. Haeffele, and R. Vidal, "Joint holographic detection and reconstruction," in *MICCAI Workshop on Machine Learning in Medical Imaging*, (2019).
11. F. Yellin, B. D. Haeffele, and R. Vidal, "Blood cell detection and counting in holographic lens-free imaging by convolutional sparse dictionary learning and coding," in *IEEE International Symposium on Biomedical Imaging*, (2017), pp. 650–653.
12. F. Yellin, B. D. Haeffele, and R. Vidal, "Multi-cell classification by convolutional dictionary learning with class proportion priors," in *IEEE Conference on Computer Vision and Pattern Recognition*, (2018), pp. 8953–8961.
13. A. Yurt, R. Stahl, G. Vanmeerbeeck, Z. Lin, M. Jayapala, and A. Lambrechts, "Towards practical cost-effective lens-free imaging," in *Optics and Biophotonics in Low-Resource Settings III*, vol. 10055 (International Society for Optics and Photonics, 2017), p. 100550J.
14. A. Anand, I. Moon, and B. Javidi, "Automated disease identification with 3-d optical imaging: a medical diagnostic tool," *Proc. IEEE* **105**(5), 924–946 (2017).
15. M.-K. Kim, "Applications of digital holography in biomedical microscopy," *J. Opt. Soc. Korea* **14**(2), 77–89 (2010).
16. S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in *Neural Information Processing Systems*, (2015), pp. 91–99.
17. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *IEEE Conference on Computer Vision and Pattern Recognition*, (2016), pp. 779–788.
18. K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in *IEEE International Conference on Computer Vision*, (2017), pp. 2961–2969.
19. B. D. Haeffele, C. Pick, Z. Lin, E. Mathieu, S. C. Ray, and R. Vidal, "An optical model of whole blood for detecting platelets in lens-free images," in *MICCAI International Workshop on Simulation and Synthesis in Medical Imaging* (Springer, 2019), pp. 140–150.
20. M. Hammer, D. Schweitzer, B. Michel, E. Thamm, and A. Kolb, "Single scattering by red blood cells," *Appl. Opt.* **37**(31), 7410–7418 (1998).
21. P. W. Kuchel and E. D. Fackerell, "Parametric-equation representation of biconcave erythrocytes," *Bull. Math. Biol.* **61**(2), 209–220 (1999).
22. B. Haeffele, R. Stahl, G. Vanmeerbeeck, and R. Vidal, "Efficient reconstruction of holographic lens-free images by sparse phase recovery," in *Medical Image Computing and Computer Assisted Intervention* (Springer, 2017), pp. 109–117.
23. J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online learning for matrix factorization and sparse coding," *The J. Mach. Learn. Res.* **11**, 19–60 (2010).